



**HAL**  
open science

## Description d'une solution logicielle, nommée Pointe au Sel, pour réaliser des analyses prototypiques

Philippe Rouet, Angélique Bègue, Pascal Mouquet

### ► To cite this version:

Philippe Rouet, Angélique Bègue, Pascal Mouquet. Description d'une solution logicielle, nommée Pointe au Sel, pour réaliser des analyses prototypiques. [Interne] Université de La Réunion; UMR 228 Espace-dev. 2021. hal-03468878

**HAL Id: hal-03468878**

<https://hal.univ-reunion.fr/hal-03468878v1>

Submitted on 14 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## RAPPORT

### Description d'une solution logicielle, nommée Pointe au Sel, pour réaliser des analyses prototypiques

## FICHE DOCUMENTAIRE

Date publication : 07/12/2021	Type de document : rapport interne
Nombre de pages :	Langue du rapport : français
Bibliographie : oui	Domaine : Sciences de l'homme et société - Sociologie
Illustrations : oui	DOI : 10.26171/ekb2-jf83
Titre du rapport : Description d'une solution logicielle nommée Pointe au Sel pour réaliser des analyses prototypiques	
Destinataire:	UMR ESPACE-DEV
Auteurs principaux: Philippe ROUET Angélique BEGUE	Organisme/Laboratoire Université de La Réunion – UMR Espace-Dev Université de La Réunion – UMR Espace-Dev
Autres collaborateurs : Pascal MOUQUET	IRD – UMR Espace Dev
Cadre de la recherche : Thèse de doctorat	
<p>Résumé :</p> <p>La théorie du noyau central, une méthode de représentation des données sociales, s'appuie sur le tri de mots exprimés et ordonnés lors d'entrevues individuelles. Ce travail fastidieux de tri a déjà été automatisé par le passé mais la solution n'est plus disponible. Ce rapport décrit une nouvelle solution logicielle permettant au chercheur de réaliser une analyse prototypique et une analyse catégorielle à partir d'un exercice d'association libre de mots, en obtenant le classement en quatre ensembles, le noyau et ses périphéries.</p>	
<p>Abstract :</p> <p>Central core theory, a social representation method, sort ranked words mentioned during individual interviews. This boring task has been computerized in the past, but the software is not downloadable anymore. This report describes a new software which allows the researcher to perform prototypical and categorical analyses from free association tasks, and obtain four sets of words, the core and his peripheries.</p>	
Mots-clés : association libre de mots ;analyse prototypique, analyse catégorielle ; représentations sociales,	
Keywords : free tasks ; prototypical rank-frequency analysis, categorical analysis; social representations	

I	CONTEXTE	4
	<b>pourquoi cet outil ?</b>	<b>4</b>
II	METHODE	4
II.1	<b>pourquoi un système de base de données relationnelle ?</b>	<b>4</b>
II.2	<b>pourquoi microsoft access ?</b>	<b>4</b>
II.3	<b>méthode conceptuelle</b>	<b>5</b>
III	FONCTIONNALITES	7
III.1	<b>Phase « initialiser »</b>	<b>7</b>
III.2	<b>phase « calibrer » le corpus</b>	<b>8</b>
III.3	<b>phase « analyser »</b>	<b>10</b>
III.4	<b>variantes</b>	<b>13</b>
III.5	<b>refaire manuellement les calculs</b>	<b>14</b>
IV	PREMIERS RETOURS D'EXPERIENCE	15
IV.1	<b>la prise en main du logiciel</b>	<b>15</b>
IV.2	<b>atouts du logiciel</b>	<b>15</b>
IV.3	<b>difficultés rencontrées</b>	<b>16</b>
V	POSSIBILITES D'AMELIORATION, D'EVOLUTION	16
V.1	<b>ameliorations</b>	<b>16</b>
V.2	<b>evolution</b>	<b>16</b>
VI	INSTALLATION DE POINTE-AU-SEL	17
VII	BIBLIOGRAPHIE	17

## I CONTEXTE

### POURQUOI CET OUTIL ?

Demande d'Angélique Bègue en mai 2020 pour une analyse statistique d'association de mots :

« Nous avons testé un exercice d'association de mots lors de mes entretiens. Il s'agissait de demander aux enquêtés de m'indiquer à partir du mot inducteur "paysage", 5 à 10 mots que ce mot leur évoquait et de les classer par ordre d'importance selon eux.

Pour analyser cet exercice, la littérature sur les représentations sociales (théorie du noyau central) indique qu'il faut calculer la fréquence d'apparition des mots, le rang moyen de ces mots (est-ce que le mot a été classé en première, deuxième, troisième position, etc). Ensuite, il faut faire une "analyse prototypique" qui consiste à croiser la fréquence d'apparition du mot et le rang moyen. Plus un mot est fréquent et plus son rang est faible (proche de 1, il a été classé en première position), plus le mot est central dans la représentation de la personne. »

## II METHODE

### II.1 POURQUOI UN SYSTEME DE BASE DE DONNEES RELATIONNELLE ?

L'outil de développement devait permettre de compter les occurrences de chaque mot d'une liste. Il devait aussi permettre la manipulation des données afin de proposer l'agrégation des différentes formes d'un mot et de ses synonymes par un tri manuel.

La solution de développement devait permettre d'aboutir à un outil exécutable sur un micro-ordinateur, si possible toute plateforme, avec une installation simple de l'environnement d'exécution, à l'exemple d'une application bureautique. L'objectif était que tout utilisateur potentiel télécharge l'outil d'une part et installe un environnement d'exécution gratuit ou très courant au téléchargement.

Les évolutions à venir de la solution de développement devront assurer une compatibilité descendante.

L'utilisation d'un système de gestion de base de données relationnelle (SGBD-R) permet la sélection, le tri, les statistiques sur des variables alphanumériques, alors qu'un tableur le permet difficilement.

### II.2 POURQUOI MICROSOFT ACCESS ?

Les SGBD-R gratuits du marché comme MySQL, PostgreSQL ne sont pas toujours simples d'installation et nécessitent un autre outil, ou un autre langage de création de l'interface homme-machine comme Java ou PHP, dont les mises à jour sont contraignantes.

Parmi les outils bureautiques, il n'y a pas de concurrence à MS Access. LibreOffice Base ne permet pas, par exemple, l'utilisation des requêtes de manipulation des données comme les insertions ou les mises à jour.

Inconvénients à ce choix : MS Access nécessite une licence payante (celle de MS Office). MS Access n'est exécutable que sur un système Windows. MS Access ne garantit pas une compatibilité des développements passés avec les futures évolutions de version d'Access, mais l'éditeur a maintenu une certaine compatibilité.

## II.3 METHODE CONCEPTUELLE

La méthode conceptuelle a porté sur trois premières phases :

- définir les données et concepts, décrire les règles de transformation des données
- définir la structure des données sous forme d'un Modèle Conceptuel des Données
- développer un prototype limité aux principales fonctionnalités : importer le jeu de données initiales, permettre à l'utilisateur de regrouper les mêmes formes d'un mot ou synonymes, afficher le résultat de l'analyse prototypique en quatre groupes ou « zones »

Le prototype a été présenté pour validation et améliorations, ce qui a permis d'ajouter à ce premier cœur logiciel stable des fonctionnalités secondaires.

La méthode s'apparente à une méthode RAD (acronyme de Rapid Application Development) dans la mesure où l'outil MS ACCESS propose une programmation par interface graphique permettant d'obtenir rapidement un prototype.

### II.3.1 REGLES DE GESTION DES DONNEES

Pour une analyse prototypique (une instance) on fournit des arguments :

Un mot, une expression permettant de distinguer le calcul d'un autre passé ou à venir.

Fichier brut : Pour n individus, on a p mots réponse qui ont un rang de 1 à p

Les réponses sont des mots ou des expressions.

« Sur une population donnée, on dispose donc d'un corpus d'items – le contenu de la représentation – et de deux indicateurs quantitatifs pour chaque élément produit : sa fréquence d'apparition et le score d'importance accordé à cet item par les sujets. » Abric JC in « Les méthodes d'étude des représentations sociales » 2003

Fréquence : nombre d'apparition de chaque mot (ou expression) = nombre de citations

Processus :

1. Calculer la fréquence intermédiaire : moyenne des fréquences
2. Pour chaque mot de la base d'analyse on calcule la somme des rangs puis le rang moyen (somme des rangs par nombre d'apparition des mots).
3. Calculer le rang intermédiaire : pour l'ensemble des items, somme des rangs (pris en compte pour le calcul) divisée par le nombre de citation de tous ces items (nombre de réponses des interviewés).

Classer les mots en quatre groupes

Rang moyen > rang intermédiaire ET fréquence > fréquence intermédiaire

Rang moyen < rang intermédiaire ET fréquence > fréquence intermédiaire

Rang moyen > rang intermédiaire ET fréquence < fréquence intermédiaire

Rang moyen < rang intermédiaire ET fréquence < fréquence intermédiaire

Options :

Pour une analyse prototypique (une instance) on fournit des arguments :

- Un rang maximum est pris en compte pour l'analyse prototypique. Ainsi entre une personne interrogée classant 4 items et une autre en classant 10, on fixe un rang maximum de validité.

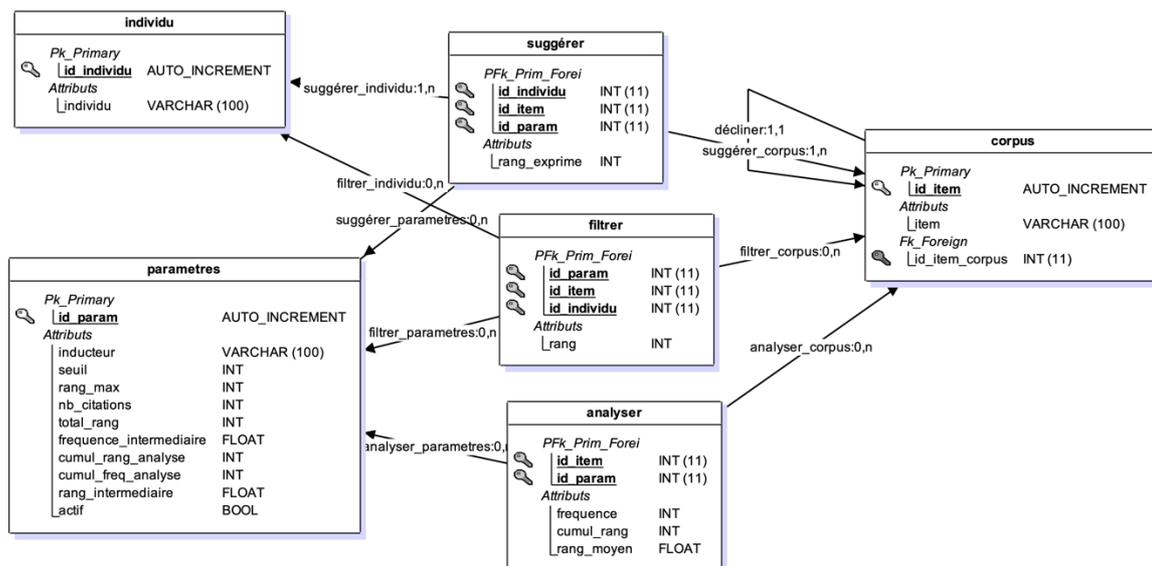
- Un nombre de citations minimal, **seuil** d'évocation en deçà duquel les mots sont exclus de l'analyse prototypique. « Le logiciel EVOC 2000 effectue ce type d'analyse sur un corpus d'évocation, pré-formaté, après un nettoyage des données qui permet de corriger les erreurs de formatage et d'apporter les corrections finales. Il calcule le rang moyen des mots qui apparaissent au moins cinq fois dans le corpus ainsi que leur fréquence d'apparition. » in Les représentations sociales des élèves du premier et du dernier degré de l'enseignement secondaire en Belgique francophone par rapport aux différentes filières.
- La « base d'analyse » consiste à filtrer le fichier brut pour avoir n individus qui ont au maximum r réponses avec un rang variant du seuil au rang maximum.
- La fréquence intermédiaire et le rang intermédiaire sont calculés sur la « base d'analyse ».

### II.3.2 STRUCTURE DES DONNEES

Le Modèle Conceptuel des Données (MCD) a été dessiné avec l'outil JMerise ([www.jfreesoft.com](http://www.jfreesoft.com)) qui permet ensuite la transformation du MCD en un Modèle Logique des données (MLD) puis un Modèle Physique des Données (MPD) dans une syntaxe SQL compréhensible et exécutable par MS Access pour créer les tables de la base de données.

Le MPD n'a été que très légèrement modifié lors du développement de l'application.

Modèle Logique des Données :



### II.3.3 TRAITEMENTS PRINCIPAUX

« L'analyse prototypique et catégorielle a pour objectif d'étudier comment s'organisent les associations libres en différenciant les éléments centraux des éléments périphériques. » (Vergès, 1992-1995)

1. Initialisation : Import du jeu de données brut, saisie d'un mot-clé et des paramètres, répartition automatique du jeu de données dans les tables.
2. Calibrer le corpus : Regroupement des déclinaisons et formes proches d'un mot  
Validation du jeu « calibré » de données
3. Classer les mots en quatre groupes :

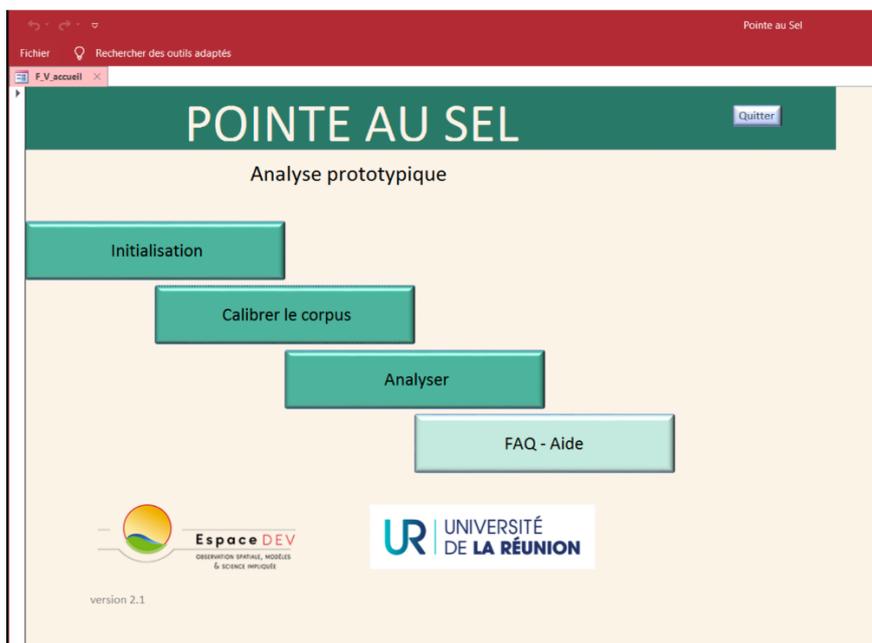
Zone 1 : fréquence > fréquence intermédiaire ET rang moyen < rang intermédiaire

Zone 2 : fréquence > fréquence intermédiaire ET rang moyen > rang intermédiaire

Zone 3 : fréquence < fréquence intermédiaire ET rang moyen < rang intermédiaire

Zone 4 : fréquence < fréquence intermédiaire ET rang moyen > rang intermédiaire

La page d'accueil du logiciel Pointe-au-Sel propose trois boutons : Initialisation, Calibrer le corpus, Analyser (un quatrième bouton propose d'accéder aux informations d'aide). Ces trois actions conduisaient aux trois traitements principaux du prototype. Par la suite, l'ajout de fonctionnalités complémentaires à ces trois fonctions font que l'action des boutons de la page d'accueil ouvre des pages à onglets.



### III FONCTIONNALITES

Le étapes internes à chaque grande fonction sont identifiées par des onglets horizontaux selon une logique de gauche à droite. Un bouton « Retour accueil » est activable.

#### III.1 PHASE « INITIALISER »

L'import concerne les données brutes telles qu'exprimées par les personnes interviewées - l'étape suivante de calibration facilitera le travail de regroupement - ou peu concerner un fichier ayant déjà été traité (calibré) avec Pointe au Sel (§ III.4.2 BOUCLE EXTERNE).

##### III.1.1 CREATION D'UN FICHER TEXTE

Créer un tableau des données, enregistré au format CSV, avec trois colonnes :

- identifiant de la personne interviewée (son nom, un code alphanumérique, numéro d'ordre...)
- mot ou expression
- rang de classement donné par interviewé

Ne pas indiquer de nom de colonne, de titre, de commentaire

##### III.1.2 IMPORT DANS UNE TABLE DE MEME STRUCTURE QUE LE FICHER TEXTE

L'import se fait par les menus de MS Access à destination d'une table interne à Pointe au Sel de même structure que le jeu de données brut et intitulée « jeu ».

##### III.1.3 CONTROLE DE LA QUALITE DE L'IMPORT

Une édition d'erreurs supposées contient les exceptions aux règles suivantes :

- Chaque item est émis par un individu au minimum

- Un individu ne classe pas deux fois le même item
- Un individu n'attribue pas deux fois le même rang

A l'usage, il est apparu que les exceptions à ces règles peuvent correspondre à une action délibérée de l'utilisateur, notamment par l'importation d'un jeu de données qui est déjà issu d'une calibration-regroupement d'items – dans Pointe au Sel.

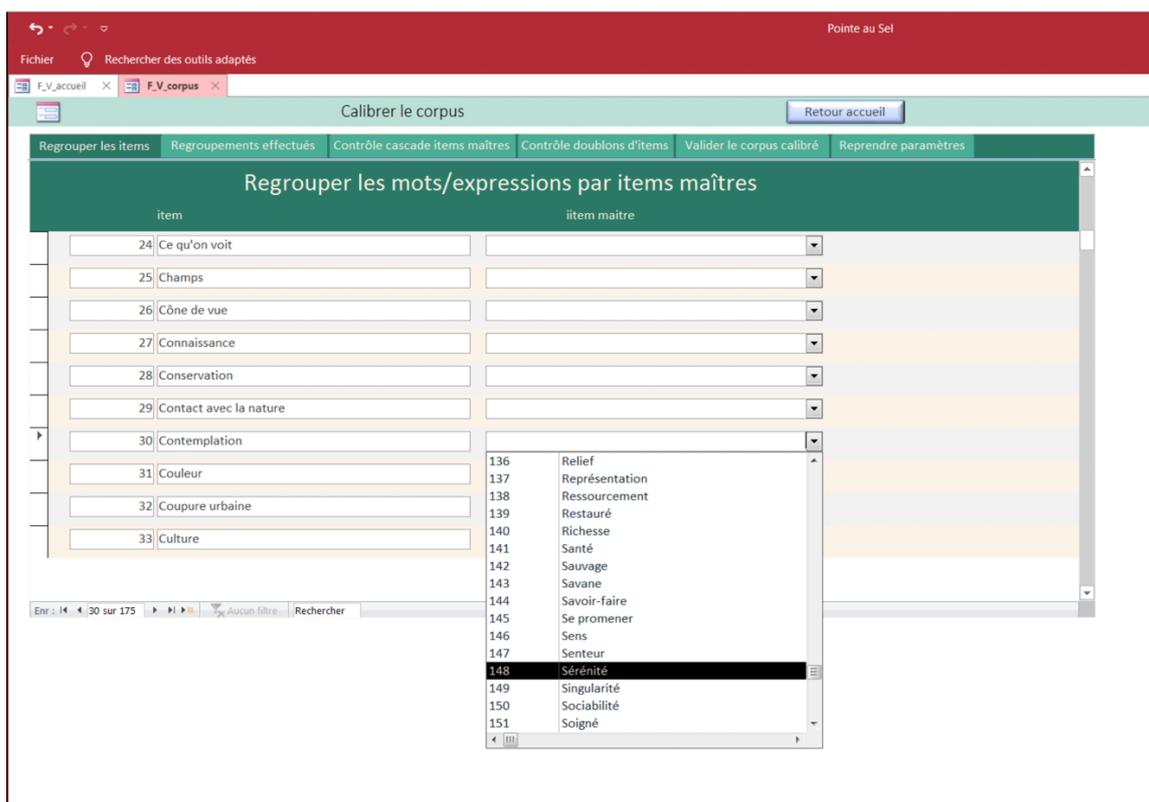
### III.1.4 INITIALISATION D'UNE NOUVELLE ANALYSE

En associant un mot-clé ou court commentaire au jeu de données, on va donner une sorte d'identifiant au jeu de données permettant à l'utilisateur de s'y retrouver après de multiples tentatives. On y ajoute deux arguments, le nombre seuil de citations d'un mot pour être inclus dans l'analyse, et le nombre de mots pris en compte par individu ou rang maximal. Le seuil d'évocation au-delà duquel les mots sont exclus de l'analyse prototypique est initialisé par défaut à 1 ; le rang maximal est initialisé par défaut à 10.

## III.2 PHASE « CALIBRER » LE CORPUS

La proposition est de permettre à l'utilisateur d'identifier parmi les items exprimés par les personnes interviewées ceux présentant une similarité pour les regrouper sous un seul item choisi comme représentatif : l'item maître. Alors que le Système de Gestion de Base de Données distingue les différentes formes d'un même mot, le pluriel ou le genre par exemple, on va laisser à l'utilisateur la maîtrise de ces regroupements ainsi que les mots ou expressions synonymes.

### III.2.1 REGROUPER LES MOTS OU EXPRESSIONS PAR ITEMS MAITRES



La liste de tous les items du fichier importé est présentée par ordre alphabétique. En regard de chaque item, une liste de valeur propose à un choix unique tous les items pour lesquels un item maître n'a pas été défini. Ces listes ne sont pas rafraîchies automatiquement mais par l'utilisateur avec la touche F9 de son clavier.

### III.2.2 CONTROLES

Trois listes permettent à l'utilisateur de faire des contrôles :

#### **Regroupements effectués**

Liste des items « feuilles » ayant un item maître défini.

#### **Contrôle cascade items maîtres**

Si les liens item - item maître sont définis en cascade, il y a une erreur.

Par exemple :

Si A a pour item maître B qui a pour item maître C

Si A a pour item maître B qui a pour item maître A

Ces erreurs doivent être impérativement corrigées.

#### **Contrôle doublons d'items**

Après regroupement, liste des items apparaissant en doublon pour une même personne interviewée.

Ces situations conduisant à une modification de l'expression des personnes interviewées concernées est laissée au choix de l'utilisateur pour la correction ou non avant l'analyse prototypique.

### III.2.3 VALIDER LE CORPUS CALIBRE



L'action sur le bouton « générer le corpus calibré » génère un jeu de données où les synonymes prennent la valeur de leur représentant, l'item maître. Ne sont pris en compte que les items dont le rang exprimé par l'interviewé est inférieur ou égal au rang max saisi en paramètre.

### III.2.4 REPREDRE LES PARAMETRES

Cette étape ne suit pas chronologiquement la calibration du corpus, il s'agit du point de départ de la variante du § III.4.1 BOUCLE INTERNE. Elle permet après une analyse menée à son terme, de remettre le dernier paramétrage à un statut « actif » pour modifier le seuil et le rang max, tout en conservant le travail de regroupement des items déjà effectué.

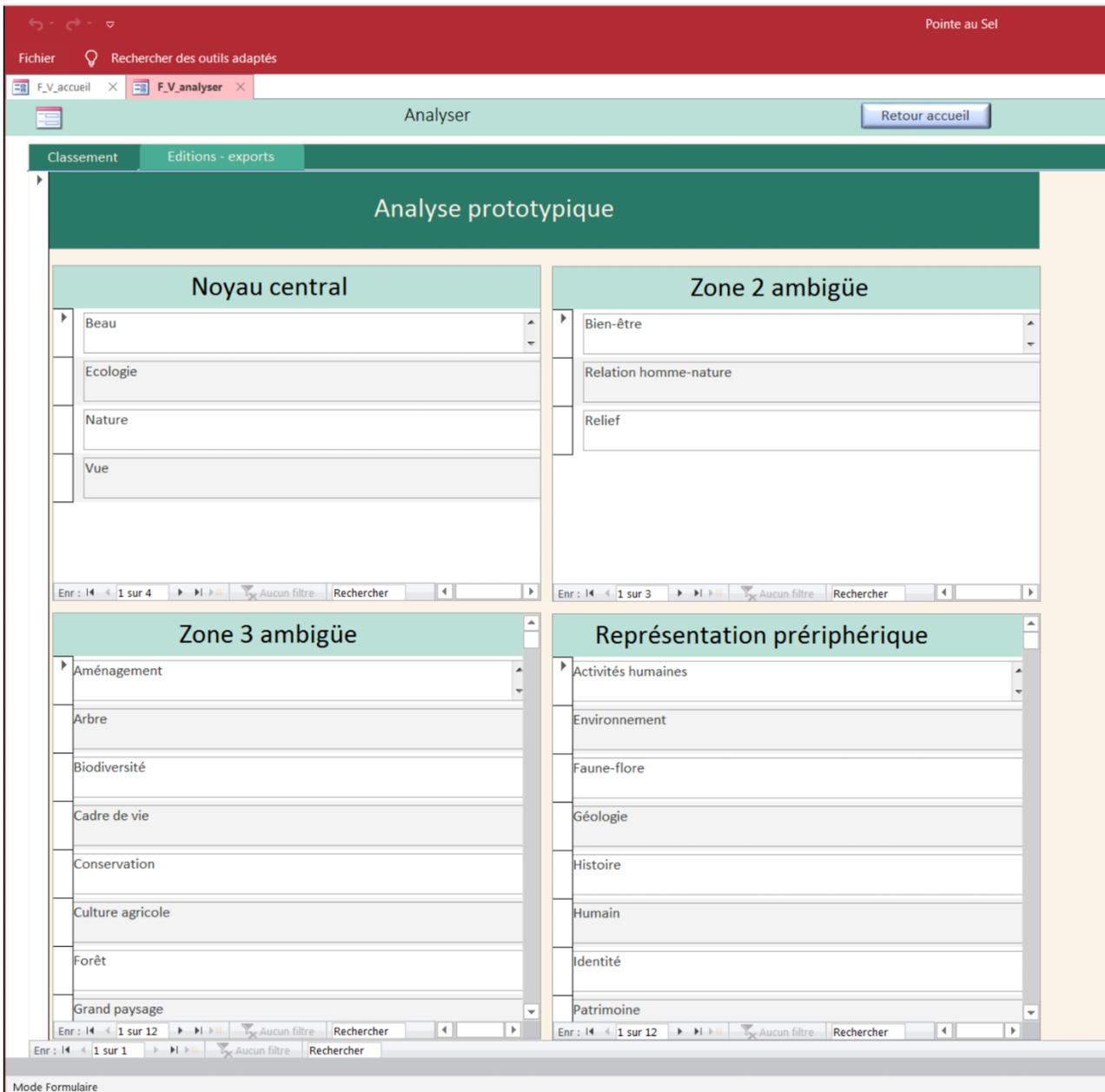
Après cette étape, le regroupement des items peut être corrigé ou maintenu, la validation du corpus calibré prendra en compte le nouveau paramètre « rang\_max ».

L'utilisateur a deux actions à effectuer:

- La réactivation du paramétrage à « actif ».
- La saisie du nouveau seuil et du rang max.

## III.3 PHASE « ANALYSER »

Cette étape va automatiquement classer et afficher les mots en quatre groupes.



Le processus va en premier calculer les résultats intermédiaires qui lui sont nécessaires.

Il va générer un jeu de données « travaillé » où le lien des items avec l'interviewé disparaît au profit du nombre d'interviewés mentionnant le mot représenté par l'item maître, la fréquence de l'item maître (ou nombre de citations par item maître).

Le jeu de données retenues va exclure les citations dont le rang exprimé sera supérieur au RANG MAX entré en paramètres

Le jeu de données retenues va être limité par la valeur de la fréquence de l'item maître qui devra égal ou dépasser le SEUIL entré en paramètre.

Le rang de l'item maître exprimé par les individus interviewés est cumulé.

Le rang moyen d'un mot est calculé avant le transfert dans le jeu d'analyse.

### III.3.1 VARIABLES DE CALCUL: FREQUENCE INTERMEDIAIRE, RANG INTERMEDIAIRE

- la valeur du nb de citations (nb de lignes de la table filtrer avant application du seuil).
- la valeur du nb d'items « maîtres » pris en compte pour le calcul de la fréquence intermédiaire

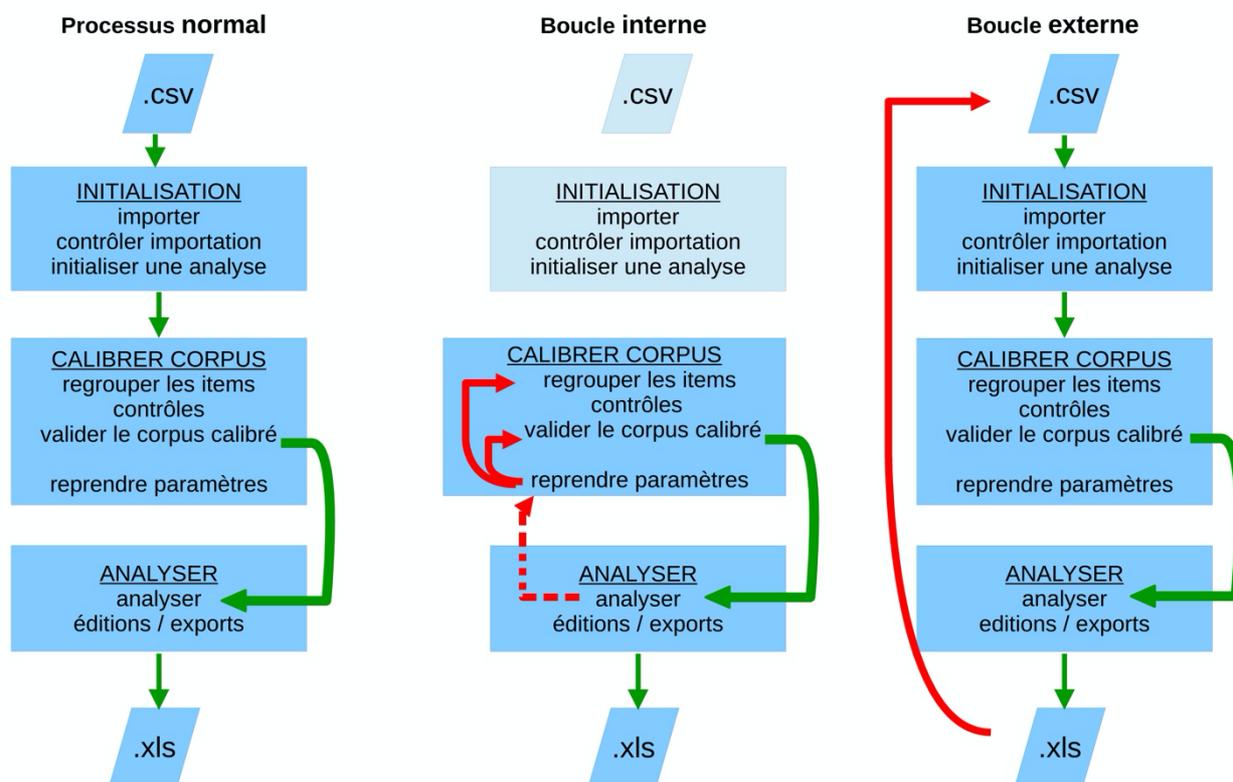
- la valeur de la fréquence intermédiaire : moyenne des fréquences de la table analyser
- La somme des rangs filtrés. N'intervient pas dans les calculs.
- La somme des rangs de tous les items pris en compte dans l'analyse (après application du seuil et du rang max)
- La somme des fréquences de tous les items pris en compte dans l'analyse (après application du seuil et du rang max)
- La valeur du rang intermédiaire : moyenne des rangs de la table analyser soit le cumul des rangs analysés sur le cumul des fréquences analysées.

### III.3.2 CLASSEMENT

Par le traitement précédent, on ne retient dans le classement que les items maîtres dont la fréquence est supérieure ou égale au seuil d'évocation et de dépassant pas le rang max saisis en paramètre. Le classement peut conduire l'utilisateur à réviser ces paramètres, c'est la variante du § III.4.1 BOUCLE INTERNE.

<p>1- Noyau central</p> <p>fréquence &gt; fréquence intermédiaire ET rang moyen &lt; rang intermédiaire</p>	<p>2- Zone ambiguë</p> <p>fréquence &gt; fréquence intermédiaire ET rang moyen &gt; rang intermédiaire</p>
<p>3- Zone ambiguë</p> <p>fréquence &lt; fréquence intermédiaire ET rang moyen &lt; rang intermédiaire</p>	<p>4- Zone de représentation périphérique</p> <p>fréquence &lt; fréquence intermédiaire ET rang moyen &gt; rang intermédiaire</p>

### III.4 VARIANTES



#### III.4.1 BOUCLE INTERNE

Outre l'ordre imposé des étapes « Initialisation – Calibrer le corpus – Analyser », après l'analyse on peut revenir à l'étape « Calibrer le corpus » en réactivant le dernier paramétrage.

Après l'action du bouton « Analyser » toute la procédure est verrouillée.

On peut reprendre la dernière analyse à partir du regroupement des items qui est conservée et la rendre modifiable.

1. Menu « Calibrer le corpus » à partir de l'accueil, puis onglet « Reprendre paramètres »  
L'état du dernier paramétrage s'affiche en fin d'une ligne d'information.  
S'il est à « Vrai » le processus n'a pas été mené jusqu'à l'analyse.  
S'il est à « Faux » :
2. Cliquer sur le bouton « Réactiver le dernier paramétrage »  
L'information « Paramétrage actif ? » est mise à jour et passe à « Vrai ».  
Les calculs intermédiaires et finaux de la dernière analyse sont supprimés. Le regroupement en items maîtres peut être modifié ou conservé.
3. Modifier si besoin les paramètres et saisir LES DEUX CHAMPS « Nouveau seuil » et « Nouveau rang max » :  
Cliquer sur le bouton « Valider nouveau seuil et rang max » : La ligne d'information sur le dernier paramétrage est mise à jour.
4. Choisir soit l'onglet « Regrouper » items, soit l'onglet « Valider le corpus Calibré »

### III.4.2 BOUCLE EXTERNE

Il est possible en fin de cycle de conserver le travail de regroupement en items maîtres effectué sur le corpus en exportant le jeu de données initial avec les « items feuilles » remplacés par les « items maîtres » sans application des paramètres « Seuil » et « Rang max ».

Ce nouveau jeu de données pourra être réimporté et analysé en suivant les étapes « Initialisation – Calibrer le corpus – Analyser ».

1. Accueil « Analyser », onglet horizontal « Éditions – exports », onglet vertical « Jeu initial avec items maîtres »
2. Exporter au format Excel
3. Transformer le fichier au format texte .csv avec le séparateur point-virgule.
4. Menu « Initialisation », onglet « Importer »

### III.5 REFAIRE MANUELLEMENT LES CALCULS

Pointe au Sel ne doit pas être une boîte noire, les résultats des étapes automatisées doivent pouvoir être vérifiés. Des éditions intermédiaires ont été prévues, une aide pour refaire les calculs à partir de ces éditions est disponible avec le logiciel.

#### III.5.1 CONTROLER L'ANALYSE PROTOTYPIQUE A PARTIR DU CORPUS ANALYSE

Rappel : pour chaque item on compare le rang moyen au rang intermédiaire et la fréquence à la fréquence intermédiaire.

A partir de l'état « corpus analysé »

Les valeurs rang moyen et de fréquence sont, par item, la valeur dans la colonne « Rang moyen » et de la colonne « Fréquence ».

fréquence intermédiaire : moyenne des fréquences du corpus analysé soit total de la colonne « fréquence » / « nb de lignes »

rang intermédiaire : moyenne des rangs du corpus analysé soit total de la colonne « cumul des rangs » / total de la colonne « fréquence »

Zone 1 : noyau central	Zone 2 : ambiguë
Zone 3 : ambiguë	Zone 4 : représentation périphérique

Zone 1 : fréquence > fréquence intermédiaire ET rang moyen < rang intermédiaire

Zone 2 : fréquence > fréquence intermédiaire ET rang moyen > rang intermédiaire

Zone 3 : fréquence < fréquence intermédiaire ET rang moyen < rang intermédiaire

Zone 4 : fréquence < fréquence intermédiaire ET rang moyen > rang intermédiaire

#### III.5.2 CONTROLER TOUTES LES ETAPES

Des listes sont éditables :

- le contenu du corpus calibré
- le contenu du corpus analysé (§III.5.1)
- la liste des items maîtres

- la liste des items exclus par un rang dépassant le rang maximum
- la liste des items dont le nombre d'évocation par les personnes interviewées est inférieur au seuil, donc ces items sont exclus de l'analyse
- la liste des items exclus par un rang dépassant le rang maximum
- le jeu de données initial avec les items feuilles remplacés par les items maîtres. Ce nouveau jeu de données n'est pas altéré par l'application du seuil et du rang maximum. Après suppression des noms de colonnes et transformation au format csv, il pourra être importé de nouveau pour une analyse prototypique ( § III.4.2 BOUCLE EXTERNE).

## IV PREMIERS RETOURS D'EXPERIENCE

### IV.1 LA PRISE EN MAIN DU LOGICIEL

Le logiciel est simple d'utilisation, il n'est composé que de trois boutons (« initialisation », « calibrer le corpus » et « analyser ») correspondant aux trois traitements principaux. Une fois qu'une étape est réalisée, il suffit de retourner à l'accueil pour procéder à la phase suivante.

Dans la phase d'initialisation, le chemin à parcourir pour importer le jeu de données est décrit de manière détaillée et les arguments à saisir pour initialiser l'analyse sont explicités dans un encadré.

Dans la phase de calibrage du corpus, un premier onglet permet de faire les regroupements de mots souhaités. Plusieurs autres onglets permettent ensuite de vérifier les regroupements effectués, contrôler les doublons, et faire apparaître les erreurs et/ou les exceptions que l'on souhaite délibérément garder avant de valider et de générer le corpus calibré.

Dans la dernière étape, le bouton analyser génère les quatre zones de l'analyse prototypique et les différents fichiers Excel éditables (III.5.2).

### IV.2 ATOUTS DU LOGICIEL

#### IV.2.1 L'EDITION DES ERREURS

Dans la phase d'initialisation, l'édition des erreurs permet de vérifier les données importées en faisant apparaître le nombre de citations d'un mot identique par un même individu, et le nombre de rang X attribué deux fois par un même individu. Les corrections sont à effectuer dans le fichier texte d'origine. Il est également possible de garder délibérément des exceptions. Dans notre cas, plusieurs mêmes individus ont classé plusieurs mots au même rang X. Ajouté à cela, l'homogénéisation du corpus a conduit à séparer des groupes de mots en plusieurs sous-entités significatives en elles-mêmes, en gardant pour chaque sous-entité le rang d'apparition du groupe de mots. Le logiciel présente ainsi l'avantage de faire apparaître les individus et les mots classés au même rang.

#### IV.2.2 LA VARIABILITE DES SEUILS

Dans la phase d'initialisation, la possibilité de faire varier les seuils de citation d'un mot et le rang max pris en compte d'une analyse à une autre, nous a permis de nous rendre compte de la valeur des seuils à choisir, en fonction de la taille de notre échantillon d'enquêtés, pour rendre à la fois la richesse et l'homogénéité des évocations au sein de l'échantillon étudié. Plusieurs tests ont été réalisés permettant de comparer le contenu des quatre zones et de constater si en fonction des différents seuils choisis ce contenu évoluait significativement ou pas.

### IV.2.3 LE CALIBRAGE DU CORPUS

Le logiciel permet de regrouper manuellement les mêmes formes d'un mot et ses synonymes, et surtout de pouvoir faire des rapprochements thématiques dans le cadre de l'analyse catégorielle. Il est donc possible de travailler à partir d'un fichier brut ou d'un fichier déjà « nettoyé ». Nous avons travaillé à partir d'un fichier déjà homogénéisé, certaines expressions récoltées pendant la phase d'entretiens devant être séparées en entités significatives en elles-mêmes puisqu'elles renvoyaient à plusieurs thématiques différentes.

Une première analyse prototypique a été réalisée nécessitant d'importer le jeu de données homogénéisé et de valider le corpus pour obtenir le classement des mots en quatre zones (III.3.1). Puis, une seconde analyse a été réalisée, avec un travail de regroupement des mots/expressions par items maîtres. Dans cette phase, nous avons procédé à des regroupements thématiques (ex : les items géologie, géomorphologie, relief ont été regroupés sous l'item maître milieu physique). Etant donné qu'un même individu pouvait avoir donné deux mots différents que nous avons décidé de classer sous le même item maître, il a été demandé de nous permettre d'avoir accès à une liste de contrôle des items en doublon pour un même individu et de prendre en compte ce cas de figure dans l'analyse. Dans cette phase, le bouton regroupement effectué a permis de contrôler et de vérifier la pertinence des regroupements effectués.

### IV.2.4 CONTROLE ET PARTAGE DES DONNEES

Dans la phase d'analyse du corpus, plusieurs fichiers Excel sont éditables. Ils permettent de vérifier manuellement les résultats mais aussi de partager avec un tiers les regroupements effectués et les résultats obtenus dans les quatre zones.

## IV.3 DIFFICULTES RENCONTREES

Dans la phase de regroupement des items, chaque item est codé. Une fois qu'un item est regroupé sous un item maître, c'est le code de l'item maître qui apparaît et non le mot en lui-même, ce qui peut rendre la tâche fastidieuse.

# V POSSIBILITES D'AMELIORATION, D'EVOLUTION

## V.1 AMELIORATIONS

Une erreur résiduelle est signalée et non résolue. Elle apparaît aléatoirement lorsque l'on veut vider les tables de la base de données pour commencer une analyse sans historique. Un bouton « vider la base de données » est prévu, il génère en cascade la suppression des lignes de toutes les tables de la base de données. Parfois, cette suppression n'est pas complète. Le logiciel « Pointe au sel » est alors bloqué, aucune importation de données n'est possible. La solution simple consiste à écraser le fichier exécutable « pointeauseau-x-y.accdr » (x et y symbolisant l'incrément de version, sous-version) par un nouveau fichier exécutable.

D'un point de vue ergonomique, la couleur des onglets est normalement changeante pour distinguer l'onglet actif, mais dans la réalité pour rafraichir les couleurs d'onglet non utilisé, l'utilisateur doit passer sa souris sur l'onglet.

## V.2 EVOLUTION

Un nouveau développement de la fonctionnalité de regroupement des items en items maîtres peut-être envisagée pour prendre en compte les premiers retours d'expérience.

## VI INSTALLATION DE POINTE-AU-SEL

Installation:

- installer Microsoft Access
- <https://framagit.org/rouet/pointe-au-sel>  
télécharger le fichier exécutable pointeausel-x-x.accdr ( 'x-x' étant le numéro de version )

## VII BIBLIOGRAPHIE

Abric, J.-C., 1994. Pratiques sociales et représentations, éd. Paris, France : PUF

Abric, J-C, 2005, Méthodes d'études des représentations sociales, éd. Toulouse, France : Erès

Ferrara, N., Friant, N., 2014, Les représentations sociales des élèves du premier et dernier degré de l'enseignement secondaire en Belgique francophone par rapport aux différentes filières, <https://doi.org/10.4000/osp.4496>

Jodelet, D., 1993. Les représentations sociales, éd. Paris, France : PUF

Moliner, P., Rateau, P., Cohen-Scali, V., 2002. Les représentations sociales. Pratique des études de terrain, Rennes, France : Presses universitaire de Rennes

Vergès, P., 1992. L'évocation de l'argent : une méthode pour la définition du noyau central d'une représentation. Bulletin de psychologie XLV, 203–209.

Vuillot, C., Mathevet, R., Sirami, C., 2020. Comparing social representations of the landscape: a methodology. E&S 25. <https://doi.org/10.5751/ES-11636-250228>